Temperature Dependent Conformational Transitions and Hydrogen Bond Dynamics of the Elastin–Like Octapeptide GVG(VPGVG): a Molecular Dynamics Study

Roger Rousseau [(1),(2)], Eduard Schreiner [(1)], Axel Kohlmeyer [(1)], and Dominik Marx [(1)]

[(1)] *Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum, 44780 Bochum, Germany*
[(2)] *International School for Advanced Studies, Via Beirut 4, 34014, Trieste, Italy*

A joint experimental / theoretical investigation of the elastin-like octapeptide GVG(VPGVG) was carried out. In this paper a comprehensive molecular dynamics study of the temperature dependent folding and unfolding of the octapeptide is presented. The current study, as well as its experimental counterpart (see previous paper) find that this peptide undergoes an "inverse temperature transition", ITT, leading to a folding at about 40-50°C. In addition, an unfolding transition is identified at unusually high temperatures approaching the boiling point of water. Due to the small size of the system two broad temperature regimes are found: the "ITT regime" (at about 10-50°C) and the "unfolding regime" at about $T > 60°C$, where the peptide has a maximum probability of being folded at $T \approx 60°C$. A detailed molecular picture involving a thermodynamic order parameter, or reaction coordinate, for this process is presented along with a time-correlation function analysis of the hydrogen bond dynamics within the peptide as well as between the peptide and solvating water molecules. Correlation with experimental evidence and ramifications on the properties of elastin are discussed.

Keywords: Elastin, molecular dynamics, inverse temperature transition, hydrogen bond dynamics.
Abbreviations used: AFM, Atomic Force Microscopy; DSC, Differential Scanning Calorimetry; CD, Circular Dichroism; FT-IR, Fourier Transform Infrared; NMR, Nuclear Magnetic Resonance; MD, Molecular Dynamics; HB, Hydrogen Bond; ITT, Inverse Temperature Transition; PCA, Principal Component Analysis; PPC, Pressure Perturbation Calorimetry.

I. INTRODUCTION

Vertebrate elastic fibers, as contained in vascular walls, skin or lungs, allow for reversible deformations upon mechanical stress. These fibers consist of two main types of protein components: a fibrous component, *fibrilen*, and an amorphous component, the globular *elastin* protein [1]. The essential elasticity is provided by the latter protein, elastin, which is known to have very unique viscoelastic properties in the water swollen state (see Refs. [2, 3, 4, 5, 6, 7] for recent overviews). The insoluble elastin is an extensively crosslinked polymer of a precursor *tropoelastin*. Tropoelastin is composed of two types of domains. One of them is rich in lysine and provides cross-linking of the individual monomers resulting in lysinonorleucin, desmosine and isodesmosine links, which are responsible for its typical yellow color. The other is made predominantly of hydrophobic amino acids with the highly repetitive pentameric repeat unit "VPGVG" (amino acids V: valine, P: proline, G: glycine) which has been found to be crucial to elastin's functionality [2, 7, 8, 9].

An interesting peculiarity of both elastin and tropoelastin, is that they undergo *folding* by *increasing* the temperature beyond typically 25°C. The term "inverse temperature transition" (ITT) was coined for this apparently paradoxical change from a "disordered" (extended) to an "ordered" (folded) conformation upon heating. A striking manifestation of this phenomena is the ability to grow crystalline solid state structures from a solution of cyclic elastin-like oligopeptides by *heating* [10, 11]. As

such, tropoelastin and its synthetic analogues, have been a subject of intense investigation in the field of biopolymers and protein engineering. Elastin-like polymers have the promise of providing materials for bio-mechanical devices [2, 8, 12] as well as temperature dependent molecular switches [9]. Despite the intense research efforts, the detailed structure of tropoelastin still defeats elucidation. Hence a molecular level picture which relates the protein structure to the viscoelastic properties is still a matter of debate. What is beyond doubt, however, is the decisive role of water as a plasticizer – dry elastin is *brittle* [13]. Still, the potentially crucial aspect of the protein's hydration water remains largely unexplored with the notable exception of recent MD simulations [14].

The origin of the viscoelastic properties of elastin is controversially discussed, calling concepts such as classical rubber elasticity [15], various librational entropy mechanism [3, 8, 16], and multi–phase models [4, 14]. Much of the experimental data has been summarized in excellent reviews [2, 6, 17] on both elastin and elastin-like polypeptides. There is much data that suggest that both elastin and its synthetic mimics display an interesting conformational dynamics in solution. Early NMR studies suggest that elastin under physiological conditions is composed of highly mobile chains [18, 19], which correlated well with the observation from birefringence [20] measurements that elastin, in the same state, is isotropically distributed with the chains adopting random conformations. Studies of the thermo-mechanical properties of water swollen elastin and tropoelastin are consistent with the interpretation of elastin as a classical

rubber with heavy emphasis on the role of entropic contributions to elasticity [15]. These findings would support a single phase model where the structure of elastin is random and undergoes large amplitude fluctuations. Recent single-molecule AFM and spectroscopic measurements [7] studies have severely challenged this interpretation and suggest that, above the ITT temperature, a structurally *ordered* model of poly(VPGVG) can also account for these elastic properties due to significant protein librational contributions to the overall entropy as suggested two decades ago [21]. Other spectroscopic experiments such as FT-IR, NMR, CD and Raman measurements [9, 22, 23, 24, 25, 26] on elastin-like polypeptides suggest the presence of β- and γ-turns as a common structural motif which may be in dynamic equilibrium and very short and/or distorted antiparallel β-strands and disordered structures with possible dynamic sheet-coil-turn transitions. Recent NMR studies also point to the high mobility of the elastin chains in the water swollen elastins [27] and in solution multiple temperature states for poly(GVGVP), each with different dynamical behavior and peptide-water interactions, were observed [28]. These later data would rather support a multiphase model of elastin [4] where the fine details of the protein structure and in particular its interactions with water play a crucial role in understanding its elasticity. Overall, there is a wealth of data on the structure and properties of these systems, and although the models do not agree on their interpretation, a central ingredient is the intriguing and complex dynamical behavior of elastin-like polymers themselves and the crucial interplay with solvating water molecules.

Atomistic simulations on elastin or elastin-analogues are scarce. Early work (see Ref. [3] for a review), such as Ref. [16], studied the elastic properties of these proteins by performing short MD simulations under constraints in order to mimic an external pulling force. Here the authors found that entropic forces from reduced librational motion of the protein were indeed a main contributing factor in the elastic behavior. However, the authors were understandably limited to simulation times not exceeding one nanosecond. It is noted that, in general, short simulation times generally lead to unreliable statistics, due to insufficient sampling of configuration space, and hence may provide only *qualitative* insights. A crucial factor also to be considered is the explicit atomistic treatment of the protein/water interface and characterization of its role in determining elastin's properties. As a significant advance along these lines, recent studies on solvated 90 amino acid polypeptides of sequence (VPGVG)$_{18}$, have been reported, which explicitly include water [14]. The importance of the protein/water interface was worked out based on thermodynamic and static structural – as opposed to dynamic – considerations. This work was successfully able to reproduce the ITT, clarify how water may contribute to the entropic contribution to the elastic forces [29] as well as probe the role of elastin mimics as potential molecular machines [30]. Overall this work

was was successfully able to qualitatively relate its findings to much of the experimentally observed phenomena, though, some debate has arisen with regards to the ability of the proposed structures to account for the temperature dependence of elastin's water content above the ITT [7]. Despite the significant improvement these simulations have brought to our understanding of the atomic level details concerning elastins' behavior, these quite long biopolymers were simulated on the order of a few nanoseconds for a given temperature. In view of the expected long relaxation times of proteins of that size [31] it would be highly desirable to have access to simulation times that allow for quantitative statistical-mechanical analysis for a given number of amino acids. This is especially true for low frequency oscillations of the protein backbone which play a central role in many models of elastin's properties. Furthermore, the *dynamics* and in particular the *kinetics* of hydrogen-bonds remains unexplored up to now. Thus, questions regarding the coupling of the dynamics of the protein, the protein/water interface and hydrogen bonding remain open issues.

A significant experimental finding was the recent demonstration [9] that oligopeptides of the kind GVG(VPGVG)$_n$ display the ITT even in the limit of only one pentameric repeat unit. Using the limiting value $n = 1$ opens up the possibility to perform MD simulations that allow for fairly long simulation times with sufficiently many solvating water molecules. We therefore have launched a joint experimental / simulation study of the smallest possible elastin model, the octamer GVG(VPGVG) (having only ≈ 640 Da), see the preceding paper [25] and Ref. [24] for a preliminary short communication. This minimal elastin model was chosen, at the expense of experimental difficulties, in order to allow for extensive molecular dynamics (MD) simulations, at 12 temperatures between 280 K–390 K, with a thorough conformational sampling of 32 nanoseconds each (after equilibration), followed by in–depth statistical mechanical analysis with a focus on dynamics. The current approach will allow us to obtain a crucial link between the multitude of experimental results and the atomistic simulations by directly comparing results obtained on essentially the same system under thermodynamically consistent conditions. Moreover, the detailed statistical mechanical analysis which is possible on the current MD simulations will provide a complementary perspective on simulation work already existing in the literature.

II. RESULTS AND DISCUSSION

A. Time evolution and Temperature Dependence of Structural Parameters

In order to get a first impression of the dynamics of the solvated octapeptide GVG(VPGVG) we consider the time evolution of a few selected quantities describing the
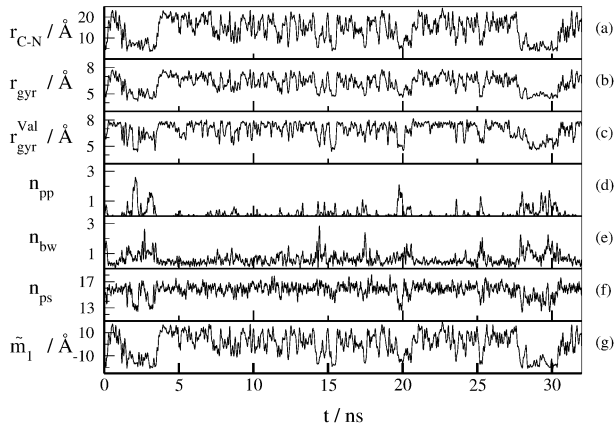
FIG. 1: Time evolution of selected structural quantities of GVG(VPGVG) in water at 330 K: (a) Distance from C– to N–terminus, r_{C-N}, (b) Radius of gyration, r_{gyr}, (c) Radius of gyration of the valine side groups, r_{gyr}^{Val}, (d) Number of peptide-peptide HBs, n_{pp}, (e) Number of waters bridging two amino acids of peptide by HBs , n_{bw}, (f) Number of HBs between peptide and solvating water, n_{ps}, (g) Projection of the trajectory onto the first eigenvector of the covariance matrix, $\tilde{m}_1$; see text for definitions. Only for presentation purposes the functions were denoised using the Savitzky-Golay procedure [32] where a time window of 128 ps and a polynomial of 6th degree was used, whereas the analysis was carried out based on the original data sets.

FIG. 2: Graphical representation of typical protein conformations at 330 K including only bridging waters. (a) Open state, (b) Closed state.

overall structure. The first parameter of interest is the distance from the C– to the N–terminus, r_{C-N}, which is a measure of the extension of this short polymer chain. A typical time evolution profile of this quantity is presented in Figure 1(a) at a temperature of 330 K. Judging from this parameter there are two distinctly different types of peptide structure: the majority of the time r_{C-N} oscillates about a distance of 17 Å (corresponding to "open structures") while a second, less probable configuration, is found when $r_{C-N} \approx 5$ Å (stemming from "closed structures"). Representative configurations, as obtained at 330 K for each type of conformation, are depicted in Figure 2. The presence of two types of peptide conformations is also mirrored by the radius of gyration of the biopolymer, r_{gyr}, which is a standard measure of a polymer's overall size. This parameter is strongly correlated with the end–to–end distance r_{C-N}, see Figure 1(b), such that for the extended structures $r_{gyr} \approx 8$ Å whereas for the closed ones $r_{gyr} \approx 4$ Å. As a simple measure of the proximity of hydrophobic side groups of the peptide we also consider the radius of gyration of the valine side groups only, r_{gyr}^{Val}, see Figure 1(c). Again, this parameter is found to be strongly, although not perfectly, correlated with the end-to-end distance and, as expected, the radius of gyration of the entire peptide.

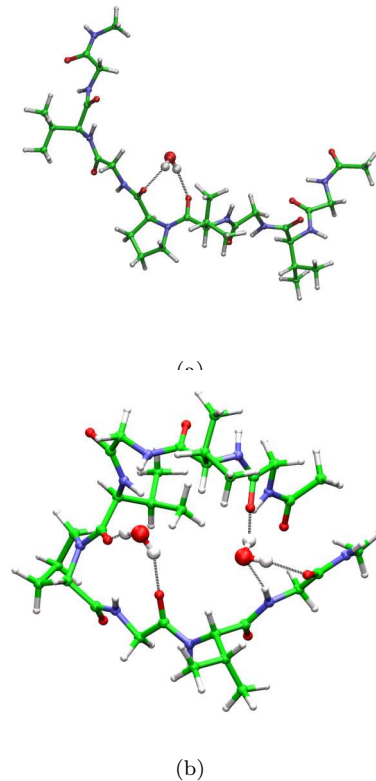The open and closed structures have distinctly dif-ferent hydrogen bond (HB) arrangements as indicated by changes in the numbers of HBs of different classes. Specifically, we consider the number of HBs within the peptide that directly connect amino acid donor (NH) and acceptor (CO) groups, n_{pp}, the number of HBs between the peptide and water molecules in its solvation shell, n_{ps}, and the number of water molecules that bridge different parts of the peptide chain by simultaneously forming HBs to two or more non-consecutive amino acid residues, n_{bw}, see Figure 2. According to Figure 1 these quantities are again found to be strongly correlated to r_{C-N}. In the closed state peptide/peptide HBs and water bridges exhibit values of around $n_{pp} \approx 2 - 3$ and $n_{bw} \approx 2$, respectively, whereas peptide/solvation water interactions amount to about $n_{ps} \approx 12$ HBs. This is quite different in the extended state: there are few peptide/peptide HBs, $n_{pp} \approx 0$, and bridging waters, $n_{bw} \approx 0$, but about $n_{ps} \approx 16$ peptide/solvation water HBs. These simple parameters – as well as others not depicted here such as the moments of inertia – indicate several opening and closing events in this trajectory and thus suggest that the present MD simulations are sufficiently able to sample the conformational space available to the peptide to distinguish between two very different types of structure which are found to be in dynamic equilibrium.

A similar picture where one observes two types of structures, open and closed, with correlated r_{C-N}, r_{gyr}

and HB types is also found in simulations where the thickness of the solvating water sphere is increased as well as with alternate peptide chain capping groups at selected temperatures of 280 and 320 K (see Appendix A 1). This gives us confidence that this scenario is not an aberration of our computational methodology. Moreover, in agreement with Ref. [9], this latter finding suggests that the results should depend on only the presence of the repeat unit VPGVG itself. We note in passing that the time scale of the peptide backbone fluctuations which interconvert open and closed conformations is on the order of about one nanosecond (see Sect. II B and in particular Figure 5(b) for details) even for this minimalistic elastin model! The implications of this time scale is that for MD trajectories on the order of only a few nanoseconds statistical quantities, which depend on the conformation of the peptide backbone, will be subject to large errors and hence should be conservatively interpreted, in particular if longer chains, yielding even longer relaxation times, are investigated.

We now consider how this dynamic equilibrium behaves as a function of temperature. Given the above description of the change in HB configuration as a function of the type of structure, it is convenient to define the number of internal hydrogen bonds, n_{IHB} as the sum of n_{pp} and the HB arising from the bridging waters, n_{bw}. Note that both n_{pp} and n_{bw} exist only when the structure is closed. An average over the entire trajectory will result in very small numbers with large associated errors, whereas, n_{IHB} is a more statistically meaningful quantity carrying similar information. The data from the time averages $< r_{C-N} >$, $< r_{gyr} >$, $< n_{IHB} >$, $< n_{ps} >$, and $< r_{gyr}^{Val} >$ are plotted in Figure 3 as a function of temperature. As expected from the above correlations $< r_{C-N} >$, $< r_{gyr} >$, $< n_{IHB} >$, and $< r_{gyr}^{Val} >$ display a similar temperature dependence which may be classified in terms of two distinct "temperature regimes" below and above a temperature close to 330 K. The average end–to–end distance increases slightly from 280 K to 310 K, where a significant decrease of $< r_{C-N} >$ sets in with a maximum contraction at 330 K, followed by a rise in value up to 390 K. Similarly, $< n_{IHB} >$ and $< r_{gyr} >$ mirror this trend. The average of the radius of gyration stemming from the valine groups only, $< r_{gyr}^{Val} >$, shows also an initial decrease, but at variance with the total radius of gyration it remains small up to high temperatures subject to sizable fluctuations. Below 330 K the structures have slightly larger r_{gyr} and only a few internal HBs, which is consistent with only a small percentage of closed structures. The largest number of closed structures is observed at 330 K resulting in the smallest $< r_{gyr} >$ and the largest $< n_{IHB} >$. Above this temperature, $< r_{gyr} >$ slowly increases in parallel to a decrease in $< n_{IHB} >$, which is consistent with a decrease in the number of closed structures. Note in passing that both $< n_{pp} >$ and $< n_{bw} >$ also follow this same trend, although with much larger numerical noise. Overall, our picture not only is consistent with the interpretation of
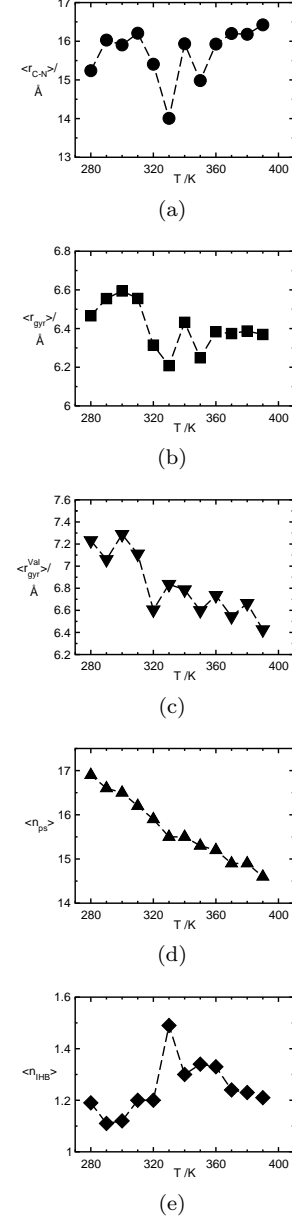

FIG. 3: Average quantities as a function of temperature: (a) Distance from C– to N–terminus, r_{C-N} (circles), (b) Radius of gyration, r_{gyr} (squares), (c) Radius of gyration of valine side groups, r_{gyr}^{Val} (triangles down), (d) Number of HBs between peptide and solvating water, n_{ps} (triangles up), (e) Number internal HBs, n_{IHB} (diamonds); see text for definitions. Dashed lines are linear connections of the data to guide the eye.

an ITT occurring at around 330 K but we also find a slow trend reversal at higher temperatures.

As a measure of how the peptide/water interface behaves during these structural transitions the total number of interfacial HBs, $< n_{ps} >$, is considered, see Figure 3(d). There is a continuous decrease in $< n_{ps} >$ observed upon increasing the temperature with a pro-

nounced kink at around 330 K. This hints that the peptide/water interface also exhibits two temperature regimes as already observed for the structural quantities associated with the peptide structure. It is noted, however, that this trend obtained from averaging the structural quantity n_{ps} is much less pronounced and is examined in greater detail in Sect. II C by dynamical analysis of the peptide/water interfacial HBs.

Overall, these averaged structural quantities provide a suggestive, though by no means comprehensive, picture of the atomic level details of the peptide structural transitions. The presence of two types of structures, open and closed, correlates extremely well with the observation of an isodichroitic point in the CD spectra [9, 25] and supports the description of this small peptide as behaving like a two-state system. The increase of the contribution of closed structures, at temperatures up to 330 K and its subsequent decrease at higher temperatures is consistent with the temperature dependence of more compact structures containing γ and β turns as observed in the FT-IR spectra of the complementary experimental investigation [25]. The DSC and PPC thermodynamic measurements presented in the joint experimental study may also be interpreted as the peptide ultimately reverting to a structurally and thermodynamically similar state at temperatures near the boiling point of water. In addition, the above scenario is qualitatively in accord with that presented in Ref. [14] which describes the ITT using a similar but much longer model peptide. Specifically this study also finds an increasing number of peptide/peptide HBs, an overall average contraction of the peptide at the ITT, a decreasing number of interfacial water molecules and an increasing number of bridging waters which may be thought of as the internal waters in our small elastin-like analogue. The authors of Ref. [14] attribute this shrinkage to a "hydrophobic collapse" which is in agreement with our observed temperature dependence of $< n_{ps} >$ and $< r_{gyr}^{Val} >$. However, the current picture suggests the subtle but crucial difference that there is a net exchange of peptide/peptide and bridging water intramolecular HBs for peptide/water interfacial HBs; the energetic implications of the various HB types are addressed in Sect. II C. At this point we stress that our finding and its interpretation has a striking resemblance to the recently advanced description of self–association of methanol in water [33] whereby the arrangement of hydrophilic–hydrophilic interactions are assumed to determine the structure as opposed to entropic hydrophobic–hydrophilic interactions. In the following two subsections these suggestive findings, which are based on simple descriptors, will be further scrutinized by a detailed statistical-mechanical analysis.

B. Principal Component Analysis and Order Parameter

To probe the above findings in a systematic but simple way one would ideally wish to describe these phenomena by a single parameter, i.e. by an order parameter or reaction coordinate. To explore this possibility we have employed what is called covariance, essential dynamics, or principal component analysis (PCA) [34, 35, 36] (see Appendix A 3 for a short introduction). In this approach one computes, by time averaging, $< \ldots >$, over the entire MD trajectory the covariance matrix $\mathbf{C} = < [\mathbf{x}(t) - < \mathbf{x} >] \cdot [\mathbf{x}(t) - < \mathbf{x} >]^T >$, where $\mathbf{x(t)}$ are the Cartesian coordinates of the peptide atoms at time t in a frame of reference where the overall translations and rotations of the polymer have been subtracted. The $3N - 6$ normalized eigenvectors $\{\mathbf{m}_i\}$ with nonzero eigenvalues $\{\lambda_i\}$ provide a basis in which the complex polypeptide motion may be decomposed; more technical details are compiled in Appendix A 3. This then provides a theoretical framework which can be though of being analogous to harmonic normal modes that are of ubiquitous use to explain spectroscopic observations of small molecules. In practice this allows us to disentangle relevant peptide backbone folding motions from small–amplitude vibrations via projecting the deviations of the MD trajectories from the average structure onto the eigenvectors, i.e. $\tilde{m}_i = [\mathbf{x}(t) - < \mathbf{x} >] \cdot \mathbf{m}_i$. Ideally, for uncorrelated small–amplitude vibrations the corresponding projections $\tilde{m}_i$ will simply fluctuate about an average value yielding an approximately Gaussian distribution function, $P(\tilde{m}_i)$; the width might be temperature–dependent akin to quasi–harmonic molecular vibrations. Strong deviations from such a unimodal behavior would single out "interesting modes" that can be associated to collective changes involving many atoms. Upon changing the temperature the various modes are expected to mix, in particular the higher–order modes. Thus, a common basis $\{\mathbf{m}_i\}$ has to be chosen to ensure a consistent analysis. In the following, projections $\{\tilde{m}_i\}$ will be performed onto the basis of eigenmodes of the low–temperature trajectory at 280 K.

Upon projecting the MD trajectories onto the eigenvector with the largest eigenvalue λ_1 it is found that this mode, $\mathbf{m}_1$, actually describes the opening and closing motion of the octapeptide, and is present at *all temperatures*, see Figure 4(a) for a graphical illustration. In Figure 1(g), we depict this projection $\tilde{m}_1$ as a function of time at 330 K where it can be inferred to be perfectly correlated with the other parameters already discussed in the previous section. The octapeptide thus exists in a closed state $\tilde{m}_1 \approx -22$ Å with end–to–end distances of about 5 Å and an extended state $\tilde{m}_1 \approx +5$ Å and end–to–end distances of about 17 Å. In addition, the projection process yields for $\tilde{m}_1$ a *bimodal* distribution function $P(\tilde{m}_1)$ at low temperatures, whereas it adopts a unimodal, but very broad and skewed shape upon heating, see Figure 4(c).
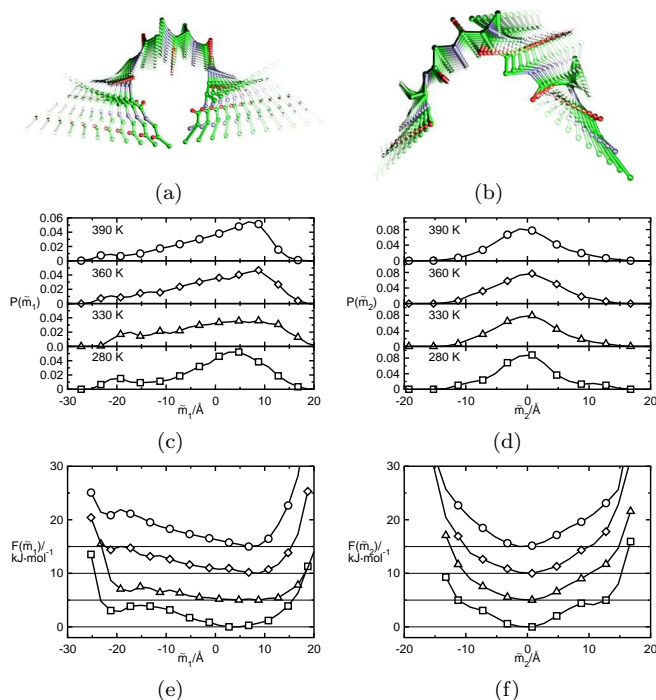
(a)   (b)

(c)   (d)

(e)   (f)

FIG. 4: Results of the Principal Component Analysis: (a) Visualization of large amplitude opening and closing mode, $\mathbf{m}_1$ (b) Visualization of large amplitude librational mode, $\mathbf{m}_2$ (c) Distribution function of the projection $\tilde{m}_1$ at 280, 330, 360 and 390 K. (d) Distribution function of the projection $\tilde{m}_2$ at 280, 330, 360 and 390 K. (e) Relative free energy along the projection $\tilde{m}_1$. (f) Relative free energy along the projection $\tilde{m}_2$. In (a) and (b) an artifical trajectory using the full first $\mathbf{m}_1$, (a), and second $\mathbf{m}_2$, (b), eigenvector was synthesized for a suitable graphical presentation. In (e)-(f) 280 K(squares), 330 K(triangles) 380 K(diamonds) and 390 K (circles). For presentation purposes the free energy profiles for 330, 380 and 390 K were shifted in energy by 5, 10 and 15 kJ·mol^{-1}, respectively.

The mode with the next largest eigenvalue, $\mathbf{m}_2$, can be classified as a peptide backbone librational mode, see Figure 4(b) for a sketch. During this motion the peptide backbone undergoes twisting which is actually very similar in spirit to that described for the deformation of β–turns in Refs. [2, 14]. However, in contrast to the lowest–order projection $\tilde{m}_1$, $\tilde{m}_2$ as well as all other higher–order mode projections of the peptide backbone are characterized by fairly symmetric, narrow and unimodal distributions $P(\tilde{m}_i)$ that are essentially temperature–independent, see Figure 4(d). This implies that – as far as the interpretation of the behavior of the folding transitions are concerned – we are able to neglect all backbone librational motions and reduce our considerations down to only a single degree of freedom: $\mathbf{m}_1$. Thus, we may confidently employ the corresponding projection $\tilde{m}_1$ as a "natural" many–body collective reaction coordinate, or order parameter, for investigating the folding transition dynamics of GVG(VPGVG) in water.

Having obtained a proper order parameter distribution

function, $P(\tilde{m}_1)$, allows one to readily define an effective relative free energy profile (or "potential of mean force") according to $\Delta F(\tilde{m}_1) = -k_{\mathrm{B}}T\ln[P(\tilde{m}_1)/P(\tilde{m}_1^{\mathrm{ref}})]$ as a function of temperature; k_{B} is the Boltzmann constant. The normalizing reference value $P(\tilde{m}_1^{\mathrm{ref}})$ is chosen such that the arbitrary value at the minimum of the free energy profile is set to zero for convenience, see Figure 4(e). At 280 K, the free energy shows two minima, separated by a small barrier, related to both open (broad global minimum at $\tilde{m}_1 \approx 5$ Å) and closed (local minimum at $\tilde{m}_1 \approx -20$ Å) structures with the latter about 3 kJ/mol higher in energy then the former. By 330 K the minimum associated with the closed structure, as well as, the barrier become much lower in relative free energy thus the entire profile of ΔF broadens out into a flat potential energy landscape. Above 330 K the free energy of the closed minimum gradually increases again and eventually disappears by 370-390 K. Performing an identical analysis for the remaining modes provides approximately harmonic potentials of mean force, see Figure 4(f) for the potential derived from $\tilde{m}_2$, in agreement with our use of only $\tilde{m}_1$ as an order parameter. The trend in $\Delta F(\tilde{m}_1)$ reflects the presence of the temperature regimes observed from the average parameters, i.e. increased folding from 280–320 K, the observation of maximum folded structures around 330 K, and the unfolding at $T \geq 340$ K. The presence of two minima in $\Delta F(\tilde{m}_1)$ in conjunction with the temperature-induced changes naturally explains the existence of the isosbestic and isodichroitic points observed in FT–IR and CD experiments [25], respectively, and gives us confidence in the reliability of our simulations.

At 330 K, fluctuations and thus statistical errors are larger then at other temperatures due to the broader configuration space which must be sampled possibly in conjunction with slower relaxation. This effect is similar to the "slowing–down problem" occurring for statistical convergence at second–order phase transitions where all quantities, and in particular the order parameter itself, undergo large fluctuations. However, the finite system size does not allow for a true phase transition with a well defined transition temperature but rather one observes, in agreement with spectroscopic and thermodynamic measurements [9, 24, 25] the two temperature regimes which are bracketed about the temperature of maximal folding, $T \approx 330$ K.

The dynamical motion of the peptide itself can be analyzed by examining the time-dependence of the reaction coordinate $\tilde{m}_1$ as described by the correlation function $c_{\mathrm{m}}(t) = \langle\tilde{m}_1(0)\tilde{m}_1(t)\rangle / \langle\tilde{m}_1(0)\tilde{m}_1(0)\rangle$, see Figure 5(a). This time auto-correlation function may be well approximated in the short time limit, $t < 100$ ps, by an exponential function, $c_{\mathrm{m}}(t) \sim \exp[-t/\tau_{\mathrm{m}}]$, with an associated relaxation time τ_{m} of the peptide's opening and closing motion. The temperature dependence of this important dynamical parameter is depicted in Figure 5(b). For temperatures below 330 K, $\tau_{\mathrm{m}} \approx 300-350$ ps is roughly constant within the accuracy of our statistics. Above 330 K,
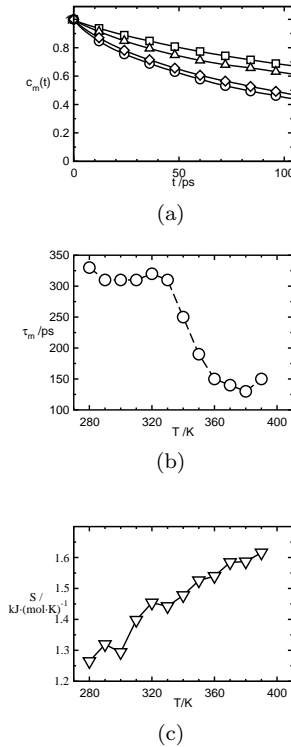
FIG. 5: Dynamics of the $\mathbf{m}_1$ eigenmode: (a) Auto-correlation function $c_m(t)$ at 280 K (squares), 330 K (triangles), 360 K (diamonds) and 390 K (circles). (b) Relaxation time τ_m as a function of temperature; here the symbol size reflects approximate error bars. (c) Peptide backbone quasi–harmonic entropy, S (down triangles), as a function of temperature; see text for definitions.

there is a progressive speed-up ($\tau_m \approx 130 - 150$ps) which reflects the steepening of the free energy surface in the region of closed structures, i.e. for $\tilde{m}_1 \ll 0$. From 370–390 K, where the free energy surface contains essentially a single minimum, $\tilde{m}_1 \approx +5$, with approximately constant curvature, the relaxation time levels off to a plateau value of about $\tau_m \approx 150$ ps. Temperature variations of τ_m thus reflect three dynamical scenarios corresponding to: *folding*, *unfolding* and *unfolded*, which is slightly different from the observations made on the average structural quantities. Note, however, that the last of the three temperature regimes represent a limiting case of the unfolding transition where there effectively no longer exists a mimimum for closed structures. These findings correlate well with larger peptide backbone fluctuations observed in NMR spectroscopy [27] for water swollen elastins. Unfortunately, the most directly comparable NMR measurements of ^{13}C relaxation times for poly(GVPGV) in solution are only obtained for the temperatures below the ITT [28]. None the less, the authors report a mean effective correlation time for polymer motion on the order of about 500 ps at 298 K in astonishing agreement with the 300 ps time scale reported here at the same temper-

ature. These facts suggest that temperature-dependent measurements of NMR ^{13}C relaxation times for this small system may serve as an independent verification of our findings.

Previous MD studies on much larger poly(GVPGV) chains find qualitatively that the peptide becomes "slightly more dynamic" or less rigid at higher temperatures [14]. This also matches our quantified speed-up of the dynamics of the peptide's backbone opening and closing motion. Thus despite the increase in number of peptide/peptide HBs, there must be an overall increase of freedom in the motion of the peptide backbone. The question then arises: does this extra motion lead to favorable entropic contribution to the ITT from the peptide? To address this, we consider the entropy change of the peptide backbone motion as estimated from a quasi-harmonic approximation based on the principal component analysis [37, 38]. Note, this analysis is based on a quasi-harmonic approximation [37, 38], whereas the lowest-order mode, $\mathbf{m}_1$, which contributes about 40% to the overall variance of the peptide backbone, is strongly anharmonic and even bimodal at low temperatures. However, its contribution to the total peptide backbone entropy is fairly small with much larger contributions arising from the more harmonic higher-order modes. Figure 5(c) shows the entropy of the peptide in a united–atom representation (see Appendix A 3) as a function of temperature. We find that despite the volume contraction of the peptide below 330 K the entropy is increasing with rising temperature, i. e. the entropy of the peptide itself *increases upon folding*. The associated entropy change is also consistent in magnitude with the experimentally measured entropy increase of 0.15 kJmol^{-1} K^{-1} and 0.11 kJmol^{-1} K^{-1} (at 298 K and 1 bar) obtained from a van't Hoff analysis of equilibrium constants of both CD and FT–IR data [9, 24, 25]. These findings are similar in spirit to various librational entropy models employed to explain viscoelastic properties of elastin and its synthetic mimetica [3, 7, 8]. Our analysis does not, however, eliminate the possibility that *stabilizing* entropic terms from the water are also important contributors to the folding behavior. Nevertheless, our simulations clearly demonstrate that our elastin mimic is a very dynamical entity and that the thermodynamic consequences of this motion, in particular the increasing entropy, are a key component in understanding its structural transitions.

### C. Hydrogen Bond Dynamics and Kinetics

In order to address the role of water and in particular the influence of hydrogen bonding on the peptide's conformational transitions we analyze the dynamics of the HB network. To this end, we employ the intermittent HB auto-correlation function [39, 40, 41, 42, 43, 44], $c(t) = \langle h(0)h(t) \rangle / \langle h \rangle$. As usual the HB population variable $h(t)$ is defined to be unity, $h(t) = 1$, if a particular

HB exists at time t and zero $h(t) = 0$ otherwise; the criteria to define a HB are compiled and discussed in Appendix A 4 and the average $\langle \ldots \rangle$ is taken over all HBs that were present at time $t = 0$. This function describes the probability that a HB, which was intact at $t = 0$, is intact at time t independently of possible breakings (and reformations) in the interim time. Similar to the average number of HBs in Sect. II A the following classes were introduced: HBs between water molecules in the bulk $c_w(t)$ (Figure 6(a)), between first shell protein solvation water and bulk water $c_{sw}(t)$ (Figure 6(b)), between the peptide and first shell solvation water $c_{ps}(t)$ (Figure 6(c)), and direct HB contacts of the peptide with itself $c_{pp}(t)$ (Figure 6(d)). These functions provides us with a wealth of insights into the complex dynamics and the associated time scales of the solvated peptide in relation to e.g. relaxation dynamics in bulk water. Comparison of $c_w(t)$ and $c_{ps}(t)$ nicely shows the well known trend [45, 46, 47] of slower dynamics for interfacial peptide water HBs then those within bulk water: depending on the temperature $c_w(t)$ and $c_{ps}(t)$ decay to zero between 10-20 ps and 20-50 ps, respectively, at temperatures below 330 K. At the highest temperatures both functions decay to zero within 10 ps with the most pronounced speed-up of this process occurring at the peptide/water interface. Conversely, the function $c_{sw}(t)$ decays to zero in only a few picoseconds at all temperatures indicating that only fast HB breaking dynamics occurs for this class of bonds. A curiosity is the function $c_{pp}(t)$, which instead of decaying to zero still retains a finite value for upwards of 100 ps or more. This behavior arises from the fact that peptide/peptide HBs, once broken cannot diffuse away from each other as they are formed involving the peptide backbone, i.e. they have a finite probability to reform even after long elapsed times due to topology.

To extract further information on the processes occurring on these various time scales we make use of a more elaborate correlation function [40, 41, 42, 43, 44, 48], $n(t) = \langle h(0)[1 - h(t)]H(t) \rangle / \langle h \rangle$, This function $n(t)$ measures the *conditional* probability that a HB, which was intact at $t = 0$, is broken at time $t > 0$ given that the donor / acceptor pair that established the HB at $t = 0$ is still close enough at time $t > 0$ to potentially form again the HB. The function $H(t)$ is unity only if the HB pair distance is smaller than a cutoff distance of 3.6 Å, which is the second nearest neighbor shell of the donor / acceptor pair radial distribution function. Thus, $n(t)$ allows one to focus on the dynamics for $t \gtrsim 2$ ps where the HB dynamics is more complex due to its coupling to translational diffusive motion [40, 41, 48], whereas the time domain for $t \lesssim 2$ ps is dominated by fast librational and possibly vibrational motion of water molecules. This function plays a significant role for two types of HBs considered in this work: HBs between water molecules in the bulk, $n_w(t)$ (Figure 7(a)), and HBs between the peptide and water molecules in its first solvation shell, $n_{ps}(t)$ (Figure 7(b)). In both cases $n(t)$ increases from $t = 0$ in view of the formation of non-hydrogen bonded pairs re-
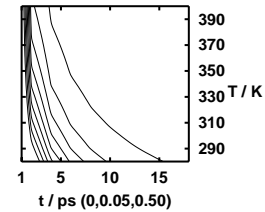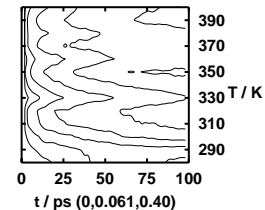

(a)
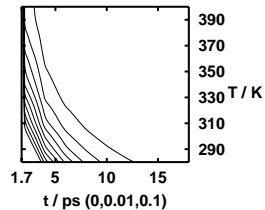
(b)
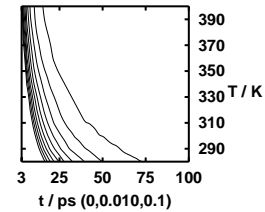
(c)

(d)

FIG. 6: Temperature and time dependence of various HB auto-correlation functions $c(t)$: HBs between (a) water molecules in the bulk c_w, (b) first shell solvation water and bulk water c_{sw}, (c) the peptide and first shell solvation water c_{ps}, and (d) direct HB contacts of the peptide with itself c_{pp}; see text for definitions. The choice of contour lines is coded in each panel as (a, b, c) where a and c denote the lowest and the highest contours, respectively, and b defines the relative spacing.
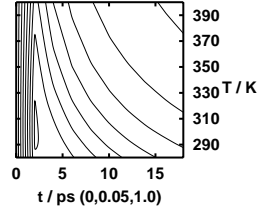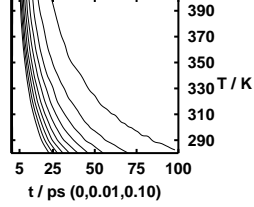
(a)

(b)

FIG. 7: Temperature and time dependence of various HB autocorrelation functions $n(t)$: HBs between (a) water molecules in the bulk n_{w} and (b) the peptide and first shell solvation water n_{ps}; see text for definitions. The choice of contour lines is coded in each panel as (a, b, c) where a and c denote the lowest and the highest contours, respectively, and b defines the relative spacing.

TABLE I: Parameters obtained from Arrhenius analysis of rate constants obtained for HB dynamics according to Eq. (1).

HB type	Symbol	T-Range	A	$E^{\ddagger}$	$\tilde{A}$	$\tilde{E}^{\ddagger}$
		K	ps^{-1}	$\mathrm{kJ \cdot mol}^{-1}$	ps^{-1}	$\mathrm{kJ \cdot mol}^{-1}$
water(bulk)-water(bulk)	w	280–390	10	7	1	5
water(bulk)-water(surface)	sw	280–390	15	7	–	–
peptide-water(surface)	ps	280–320	60	12	17	10
peptide-water(surface)	ps	320–390	2	4	0.2	≈ 0
peptide-peptide	pp	280–320	30	8	–	–
peptide-peptide	pp	330–390	4	6	–	–

some of this discrepancy may arise from systematic errors as discussed in Appendix A 4, it is more likely due to the difference in the dynamical properties of TIP3P water from SPC or SPC/E models (see Ref. [49] and references therein) used in previous studies [40, 41, 44, 48]. In particular, TIP3P water has a larger diffusion coefficient [49] which allows the water molecules to diffuse away at a different rate after an HB breaking event.

Although the bulk water behaves in an expected manner the same cannot be said for HBs that connect the peptide to its surrounding water shell. As with the average structural parameters and the peptide backbone relaxation time, both k_{ps} and $\tilde{k}_{\mathrm{ps}}$ show *two distinct* Arrhenius temperature regimes bracketed around 330 K. Below 320 K the activation energy $E_{\mathrm{ps}}^{\ddagger} \approx 12$ kJ/mol to break an interfacial HB is higher than that in the bulk, whereas above 330 K it is much lower (≈ 4 kJ/mol). Concurrently, HB reforming at the interface: $\tilde{E}_{\mathrm{ps}}^{\ddagger} \approx 11$ kJ/mol is close to that for breaking, but at 330 K the apparent activation energy ($\tilde{E}_{\mathrm{ps}}^{\ddagger} \approx 0$ kJ/mol) vanishes. In terms of rates, see Figure 8, both processes to break and reform HBs at the peptide/water interface effectively *slow down upon heating* (w.r.t. a linear Arrhenius–extrapolation of the low temperature behavior) once a critical threshold, about 320–330 K, is surmounted. Hence, the rate to make interfacial HBs essentially reaches a plateau value above 330 K, whereas the one to break these bonds continues to grow with temperature, see Figure 8.

This observation of a profound change in the *dynamics* of the peptide/water interface is in qualitative accord with thermodynamic measurements [7, 24, 25]. In particular, the behavior of the thermal expansion coefficient also indicates an overall decrease in the "strength" of these interactions. This is also consistent with the proposed hydrophobic collapse which occurs for the ITT of elastins [14]. However, the most intriguing aspect of this observation is that the peptide/water HB dynamics, which occurs on a time scale of 1-10 ps, mirrors the dynamics of the peptide which occurs on time scales that are two orders of magnitude larger.

We now investigate the dynamics of the HBs between interfacial water and bulk water as well as those involving the bridging waters. For the latter class a correlation function $c_{\mathrm{bw}}(t)$ is introduced based on a HB population function $h_{\mathrm{bw}}(t)$ which is unity only if a water molecule is *simultaneously* bound to two or more different residues

sulting from initial HB *breaking*, up to a maximum point and then decreases steadily to zero as the probability of *reforming* the HB decreases due to diffusion. Comparison of Figure 7(b)(a) and (b) shows a more rapid decay for $n_{\mathrm{w}}(t)$ than $n_{\mathrm{ps}}(t)$, which is due to the slower dynamics of HBs at the peptide/water interface.

Following previous studies [40, 41, 48], we formulate a rate equation describing the kinetics of HBs as a combination of terms resulting from *breaking* and *reforming* HBs

$$-\frac{dc(t)}{dt} = kc(t) - \tilde{k}n(t) , \qquad (1)$$

where k is the rate constant of HB breaking and $\tilde{k}$ is that for HB reformation subsequent to breaking; note that $\tau = 1/k$ defines an average HB lifetime. For bulk water, see Figure 8(a) and Table I, both time constants k_{w} and $\tilde{k}_{\mathrm{w}}$ show a simple Arrhenius or classical transition–state–theory behavior, i.e. $k = A \exp[-E^{\ddagger}/k_{\mathrm{B}}T]$, at all temperatures (with $E_{\mathrm{w}}^{\ddagger} \approx 7$ kJ/mol and $\tilde{E}_{\mathrm{w}}^{\ddagger} \approx 5$ kJ/mol). Thus, both HB breaking and reformation increase in rate *constantly* with temperature in the bulk, as expected for a simple thermal process. Note that the bulk water value of k_{w} at 300 K is close to data from simulations of pure water at room temperature [40, 41, 44, 48] whereas our value of $\tilde{k}_{\mathrm{w}}$ is smaller by a factor of about one half. Although
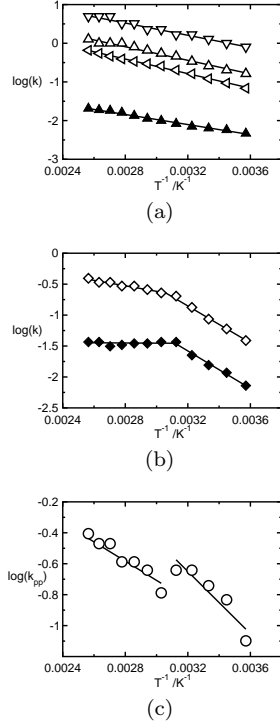
FIG. 8: Arrhenius plot of rate constants for HB breaking, k, and reformation, $\tilde{k}$, for HBs between (a) first shell protein solvation water and bulk water k_{sw} (down triangles), water molecules in the bulk k_w (open up–triangles) and $\tilde{k}_w$ (filled up–triangles); (b) the peptide and first shell solvation water k_{ps} (open diamonds) and $\tilde{k}_{ps}$ (filled diamonds); (c) the peptide with itself k_{pp} (circles). The symbol size covers the error bars and the lines are linear (i.e. Arrhenius) fits; see text for definitions and Table I for the resulting fit parameters.

of the peptide. For the former of these two HB classes the second term in Eq. (1) may be neglected due to the relatively rapid decay of $c_{sw}(t)$ to zero. We find that k_{sw} features essentially an Arrhenius behavior at all temperatures with a similar value of $E^{\ddagger}_{sw} = 7$ kJ/mol as bulk water, see Figure 8(a) and Table I. This picture is entirely in accord with the proposition that these surface water molecules have a hindered orientation which strains the HBs that are formed with bulk water. This does not allow for HB reformation after breaking, hence the process of fast HB reformation does not occur for this class of HBs. More surprising is that the correlation function of the bridging waters, $c_{bw}(t)$, essentially decays to zero within 2 ps at *all* temperatures (not shown for that reason). For this small peptide it may thus be concluded from the *dynamics* that, although water–mediated peptide/peptide bridges ("internal water") do occur, they play no key stabilizing role for the folded state. As a corollary for the above assertion of interchange of HBs to be at the heart of the folding / unfolding behavior of elastin, the stability of the folded state must then come from the peptide/peptide HBs.

We thus turn our focus to the function c_{pp} which de-

scribes the dynamics of the peptide/peptide HBs. Unfortunately, the poor statistics resulting from the few HBs that exist for $T < 320$ K, and the relatively poor convergence for long-time decay only allow for a partial analysis. Nonetheless, we are able to make several important observations. First, the short-time regime $t < 4$ ps can be used to obtain an approximation to k_{pp} in a similar fashion as above, see Figure 8. Below 330 K, k_{pp} may be interpreted as having approximately Arrhenius behavior with $E^{\ddagger}_{pp} \approx 8$ kJ/mol, see Figure 8(c). This barrier is similar to that for HBs between bulk water but less then those between the peptide and solvent waters; see Table I for a detailed comparison. Above 330 K, where a larger number of peptide/peptide HBs improves the statistics of $c_{pp}(t)$, k_{pp} is again found to have an approximate Arrhenius behavior. At these temperatures the peptide/peptide HBs exhibit a slightly lower activation barrier for breaking of $E^{\ddagger}_{pp} \approx 6$ kJ/mol, which is *larger* than that for peptide/water HBs in the same temperature regime, and will thus *stabilize the folded state*. The increase of thermal energy which allows peptide/water HBs to break more easily decreases the stability of the open state but has less effect on the folded state which is in part stabilized by peptide/peptide HBs. Second, although the long time decay of the function has too much numerical noise to provide a reliable fitting it is evident that the probability to recover the same HB pair decreases rapidly with increasing temperature above 330 K, see Figure 6(d). This is in accord with the above observation of a speed-up of the peptide's backbone motion in this same temperature regime. Increasing fluctuations of the peptide backbone above 330 K result in breaking peptide/peptide bonds more rapidly and decreasing the probability with which they reform. This effect will counter the stabilization from peptide/peptide HBs for the folded state and will eventually lead to the unfolding at higher temperature.

III. CONCLUSIONS

The presented study is successfully able to reproduce the inverse temperature transition of a minimal elastin model at about 40-50°C. Additionally, an unfolding transition is identified at temperatures approaching the boiling point of water. Our simulations reproduce the key observations obtained from FT–IR, CD, DSC, and PPC experiments in the companion study [25] of the same water solvated octapeptide. Due to the small size of the system two broad temperature regimes are found both in experiment and in the simulations: the "ITT regime" (at about 10-50°C) and the "unfolding regime" at about $T > 60$°C where the structure has a maximum probability of being folded at $T \approx 50$°C. A detailed molecular picture involving a reaction coordinate and a free energy profile for this process is presented along with a thorough time-correlation function analysis of the hydrogen bond dynamics and kinetics within the peptide as well as at the

peptide/water interface. In particular, the two regimes are characterized by changes in a dynamical two-state equilibrium between open and closed conformations and a change of the HB dynamics involving interfacial water molecules.

In a nutshell, our data suggest a simple *qualitative* picture of the observed transitions. At low temperatures, a relatively strong peptide/water interaction stabilizes open conformations relative to closed ones. The increase of thermal energy however decreases the stability of the extended state but has less effect on the folded conformation which is in part stabilized by peptide/peptide HBs. Thus, there is a shift in equilibrium and one observes an increase in folded structures and hence the ITT. A second important contributing factor is the increase of peptide backbone fluctuations above the ITT and the resulting entropic stability. However, the large librational amplitude motion of the peptide backbone can provide sufficient thermal excitation to break these peptide/peptide HBs at higher temperatures which ultimately leads to a second, *unfolding*, transition.

The current work, based upon a temperature scan using 32 ns simulations of an octapeptide, gives further support to the conclusions of previous studies employing much shorter MD simulations on longer poly(VPGVG) peptides [14, 17, 29, 30]. As in these studies we also observe a decrease in the peptide/water interactions around the ITT and a speed-up of the peptide backbone dynamics above this transition. Although our longer simulation times do *qualitatively* agree with the findings of this earlier work the present study is able to provide a more *quantitative* analysis of the *dynamical processes* associated with the peptide structural transitions. Novel findings related to the inter-relation of of peptide/water HB dynamics and peptide librational entropy, which are complementary aspects of the ITT and unfolding transitions as opposed to being alternate models, are only possible due to these quite long MD trajectories. Thus, the current work emphasizes the importance of detailed statistical-mechanical analysis in constructing atomic level pictures of complex biomolecular processes.

The limitation of the current approach however, is that it is not at present clear how all of the factors, which have been isolated as key ingredients in the peptide structural transitions, will apply to large polypeptides. Although it is highly likely that decrease in peptide/water HB interactions and peptide librational entropy will remain important ingredients the relative weights of the two factors in larger polymers is unknown. Most notably,the energetic/entropic contribution of bridging waters in larger polymers where the folded peptide may be truly said to have an "inside" is currently an open issue. To address these questions, future combined experimental and theoretical work on GVG(VPGVG)$_n$ with $n = 2$ and 3 including mutant structures as well as tropoelastins is currently underway.

APPENDIX A: THEORETICAL METHODS

1. System

The system consists of an eight amino acid oligopeptide of sequence GVG(VPGVG) capped by methylamine (–NH-CH$_3$) and acetyl (–CO-CH$_3$) groups at its C– and N–termini, respectively. In accordance with the experimental finding [9] that VPGVG is the minimal repeat unit necessary for observing the ITT irrespective of the end groups, the caps are not expected to change the basic scenario as found in our complementary experimental study on the zwitter ionic species [24, 25]; further checks are discussed below. The octapeptide is solvated with 2127 waters and its center of mass is fixed to the center of a 50 Å diameter spherical droplet surrounded by stochastic boundary conditions [50]. This sphere diameter allows us to maintain at least three to four molecules between the peptide and the stochastic boundary even for the configurations where elastin is at its maximum elongation of about 27 Å. We employ the all atom CHARMM force field [51] for the peptide and the TIP3P [52] water potential.

2. Simulations

Within the EGO molecular dynamics package [53, 54] a weak Berendsen thermostatting scheme (coupling time constant of 100 fs) was used to control the temperature using a 1 fs MD time step and the SHAKE algorithm for fixing the bonds between hydrogen and heavy atoms. This conservative choice of parameters for a careful integration of the equations of motion has allowed us to perform stable simulations which included equilibration times of 2–5 ns (depending on temperature) prior to obtaining 32 ns trajectories at 12 temperatures between 280 and 390 K. An additional run was performed at 400 K (not shown) in order to stretch the temperature range as much as possible beyond the boiling point of water. For many properties, such as HB breaking and reformation rate constants, the same trend as obtained from "extrapolation" of the data from 340 to 390 K was observed. However, in particular average structural parameters showed deviations from this trend which we interpret tentatively as failures of the simulation approach.

Note that 400 K exceeds by about 100 K the temperature range for which the employed force fields for both water and the peptide are optimized, i.e. ambient conditions. As a convergence check an additional 8 ns run was obtained at 280 K in order to test the reliability of both average and dynamical properties of the protein; the results at 280 K are quoted for the 32 ns MD run for statistical consistency. Internal pressures inside the droplet were found to be less than about 0.1–0.2 kbar depending on the temperature, which is well below the pressure of the order of many kbars that is required to suppress protein folding entirely [25, 55].

In order to cross–check the possible influence of the capping groups we have computed two 25 ns trajectories at 280 and 320 K in larger droplets (60 Å diameter with 3700 water molecules) for both the capped and zwitterionic peptides. We obtained qualitatively identical results to those presented here indicating that the ITT is not significantly affected by the presence of the caps, which is in line with experimental findings [9], nor by the spherical boundary conditions for the chosen size of the solvating droplet.

3. Principle component analysis

Principal component analysis (PCA) is a standard projection method in statistical data analysis, feature extraction, and data compression [36]. Given a set of multivariate data, the purpose is to find a smaller set of variables with less redundancy reproducing the original data as well as possible. The redundancy is measured by correlations between the data elements.

The starting point in PCA is a random vector $\mathbf{x}$ with n elements and a sample $\mathbf{x}^{(1)}$, ..., $\mathbf{x}^{(L)}$ from this vector. No assumptions on a generative model or the probability density of the vectors are made as long as the first- and second-order statistics can be estimated from the sample.

First of all the vector $\mathbf{x}$ is centered

$$\mathbf{x} \leftarrow \mathbf{x} - <\mathbf{x}> \qquad (A1)$$

and the mean $<\mathbf{x}>$ is estimated from the available sample. Geometrically this transformation corresponds to a translation of the entire sample so that the mean is now at the origin. Next consider a linear combination

$$\tilde{m}_1 = \sum_{k=1}^{n} w_{1_k} x_k = \mathbf{w}_1^T \mathbf{x} \ , \qquad (A2)$$

where the index goes over all elements of $\mathbf{x}$. The weights $w_{1_1}, \ldots, w_{1_k}$ are real numbers, elements of an n-dimensional vector $\mathbf{w}_1$ and $\mathbf{w}_1^T$ denotes the transpose of $\mathbf{w}_1$. The factor $\tilde{m}_1$ is called the first principal component of $\mathbf{x}$, if the variance of $\tilde{m}_1$ is maximally large. Therefore one has to maximize the second moment

$$<\tilde{m}_1^2> = <(\mathbf{w}_1^T \mathbf{x})^2> = \mathbf{w}_1^T <\mathbf{x}\mathbf{x}^T> \mathbf{w_1} = \mathbf{w}_1^T \mathbf{C_x} \mathbf{w}_1 \qquad (A3)$$

subject to the normalization condition

$$|\mathbf{w}_1| = (\mathbf{w}_1^T \mathbf{w}_1)^{1/2} = 1 \ , \qquad (A4)$$

where $|\cdots|$ denotes the euclidian norm

$$|\mathbf{w}_1| = (\mathbf{w}_1^T \mathbf{w}_1)^{1/2} \qquad (A5)$$

of $\mathbf{w}_1$ and $\mathbf{C_x}$ is the $n \times n$ covariance matrix of $\mathbf{x}$. The maximization problem can be rewritten

$$(\mathbf{C_x} - \lambda \mathbf{I})\mathbf{w}_1 = \mathbf{0} \qquad (A6)$$

or

$$\det(\mathbf{C_x} - \lambda \mathbf{I}) = \mathbf{0} \qquad (A7)$$

which means that λ is an eigenvalue of $\mathbf{C_x}$ and that $\mathbf{w}_1$ is the corresponding eigenvector

$$\mathbf{w}_1 = \mathbf{m}_1 \ . \qquad (A8)$$

Therefore the first principal component is a projection of the vector $\mathbf{x}$ onto the unit length eigenvector of it's covariance matrix having the largest variance, i.e. eigenvalue.

This solution can be generalized to n principal components where the i-th principal component is

$$\tilde{m}_i = \sum_{k=1}^{n} m_{i_k} x_k = \mathbf{m}_i^T \mathbf{x} \qquad (A9)$$

with $1 \leq i \leq n$ and the eigenvectors are ordered in descending order of the magnitude of their eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. Thus the original data set can be fully reproduced by the expansion

$$\sum_{i=1}^{n} \tilde{m}_i \mathbf{m}_i = \sum_{i=1}^{n} \mathbf{m}_i^T \mathbf{x} \ \mathbf{m}_i \ . \qquad (A10)$$

The hope is now that by considering just the first few principal components most of the data can be represented as

$$\mathbf{x} = \sum_{i=1}^{n} \tilde{m}_i \mathbf{m}_i \approx \sum_{i=1}^{j<n} \tilde{m}_i \mathbf{m}_i = \mathbf{x}_{\text{approx}} \qquad \text{i.e. } \mathbf{x}_{\text{approx}} \approx \mathbf{x} \ . \qquad (A11)$$

In the case of analyzing MD trajectories the variable $\mathbf{x}$ can be an $3N$-dimensional vector consisting of the cartesian coordinates of the N atoms under consideration in the form

$$\mathbf{x} = (x_1, y_1, z_1, x_2, y_2, z_2, \cdots, x_N, y_N, z_N) \qquad (A12)$$

and the sample $\mathbf{x}^{(1)}$, ..., $\mathbf{x}^{(L)}$ is given by a MD trajectory of finite length L, see e.g. Refs. [34, 35]. Before performing the PCA itself the data set must be corrected for overall translation and rotation. After centering according to Eq. A1 the overall rotation of the considered system is

subtracted. This can be done by rotating each configuration to a reference configuration by a least squares fit [56]. This translation-rotation correction will thus provide six eigenvalues equal to zero, which do not refer to the purely internal motion.

Since the eigenvectors of the covariance matrix are mutually orthogonal, i.e. uncorrelated, the motions in the direction of each component, i.e. along the corresponding eigenvector, can be interpreted as a "mode". To visualize the motion of each mode one has to calculate the coordinates including the mean of $\mathbf{x}$ and the fluctuations of the positions described by the principal component under consideration

$$\mathbf{x}_i^{(l)} = \tilde{m}_i^{(l)} \mathbf{m}_i + <\mathbf{x}> \qquad (A13)$$

where the index i denotes the i-th principal component and l the l-th timestep of the trajectory of total length L. Note that this yields a "caricature" of an important dynamical motif hidden in the full dynamics [34].

For the PCA carried out here, all heavy atoms of the peptide backbone were included. In addition, a united–atom representation was introduced for each side chain by assigning a mass equivalent to that of all its constituting atoms which was located at its center of mass. This procedure effectively reduces the number of atoms to $N = 38$ and thus simplifies the analysis and its representation. In line with previous findings [57] it was confirmed that this approximation does not lead to any significant alteration of the results. Tests indicate that PCA including all atoms provides an essentially identical picture for the low–order modes that are of interest. The dependence of the resulting motions upon trajectory length was carefully evaluated and results were found to be well converged for simulations on the order of 20 ns in duration. For example, the dot product between the slowest motion, i.e. $\{\mathbf{m}_1\}$ computed over two distinct 20 ns trajectories and that from a 30-40 ns trajectory at the lowest temperature, 280 K, is about 0.98 therefore indicating that the modes are essentially converged. Finally, the modes discussed in this investigation are found to exist also when using the zwitterionic octapeptide and in the simulations with a larger solvation sphere around the peptide as discussed in Appendix A 2.

4. Hydrogen Bond Analysis

Concerning the definition of HBs in water, and thus the population variable $h(t)$, we employ a standard structural criterion based on both the distance ($R_{\mathrm{OH}} < 2.6$ Å as obtained from including the entire first nearest neighbor peak in the intermolecular OH radial distribution function) and the angle, $\angle_{\mathrm{O-H\cdots O}} > 130°$ for $h(t) = 1$ [39, 40, 41, 42, 44, 48]. As a check we also performed the analysis with only the distance criterion and yet again with the distance criterion but using a tighter angular criterion of $\angle_{\mathrm{O-H\cdots O}} > 150°$. Consistent with previous findings [42, 44] the *absolute* value of rate constants, k_w

and $\tilde{k}_w$, is sensitive to the actual choice of the parameters of the definition and was found to vary as much as a factor of 2-3 between different definitions. However, we always obtained the same qualitative trends that both HB numbers and rate constants (k_w and $\tilde{k}_w$) feature a linear behavior with respect to temperature. Similarly, for peptide/peptide and peptide/water HBs the same criterion based on distances ($R_{\mathrm{AH}} < 2.6$ Å where the proton acceptor A is either an O atom of a water molecule or of a peptide carbonyl group) and angles ($\angle_{\mathrm{D-H\cdots A}} > 130°$ with the proton donor D being either water O or peptide backbone N) was used. As a check we varied the distance to $R_{\mathrm{AH}} < 2.4$ Å keeping the angle cutoff unaltered and performed the analysis with only the distance criterion. Again we found the same general behavior of two temperature regimes based on the dynamics of peptide/water HBs and the temperature dependence of n_{pp} and n_{bw} to be preserved. In conclusion, although the HB definitions do change the reported *absolute values* of the HB numbers and rate constants as expected [42, 44], the qualitative behaviour remains the same and thus the trends in the temperature dependence for these quantities are not affected.

[1] Pasquali-Ronchetti, I., Fornieri, C., Baccarani-Contri, M., & Quaglino, D. (1995). Ultrastructure of elastin. In: *The molecular biology and pathology of elastic tissues* pp. 31–42, Ciba Foundation : John Wiley & Sons Ltd.

[2] Urry, D. (1997). Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *Journal of Physical Chemistry B,* **101**, 11007–11028.

[3] Debelle, L. & Tamburro, A. (1999). Elastin: molecular description and function. *The International Journal of Biochemistry & Cell Biology,* **31**, 261–272.

[4] Debelle, L. & Alix, A. (1999). The structures of elastins and their function. *Biochimie,* **81**, 981–994.

[5] Reiersen, H. & Rees, A. (2001). The hunchback and its neighbours: proline as an environmental modulator. *Trends in Biochemical Sciences,* **26**, 679–684.

[6] Martino, M., Perri, T., & Tamburro, A. (2002). Biopolymers and biomaterials based on elastomeric proteins. *Molecular Bioscience,* **2**, 319–328.

[7] Urry, D., Hugel, T., Seitz, M., Gaub, H., Sheiba, L., Dea, J., Xu, J., & Parker, T. (2002). Elastin: a representative ideal protein elastomer. *Phil. Trans. R. Soc. Lond. B,* **357**, 169–184. doi:10.1098/rstb.2001.1023.

[8] Urry, D. (1993). Molecular machines: how motion and other functions of living organisms can result from reversible chemical changes. *Angewandte Chemie International Edition In English,* **32**, 819–841.

[9] Reiersen, H., Clarke, A., & Rees, A. (1998). Short elastin–like peptides exhibit the same temperature–induced structural transitions as elastin polymers: implications for protein engineering. *Journal of Molecular Biology,* **283**, 255–264.

[10] Urry, D., Long, M., & Sugano, H. (1978). Cyclic analog of elastin polyhexapeptide exhibits an inverse temperature transition leading to crystallization. *Journal of Biological Chemistry,* **253**, 6301–6302.

[11] Cook, W., Einspahr, H., Trapane, T., Urry, D., & Bugg, C. (1980). Crystal structure and conformation of the cyclic trimer of a repeat pentapeptide of elastin, cyclo-(L-valyl-L-prolylglycyl-L-valylglycyl)$_3$. *Journal of American Chemical Society,* **102**, 5502–5505.

[12] Nath, N. & Chilkoti, A. (2001). Interfacial phase transition of an environmentally responsive elastin biopolymer adsorbed on functionalized gold nanoparticles studied by colloidal surface plasmon resonance. *Journal of American Chemical Society,* **123**, 8197–8202. doi:10.1021/ja015585r.

[13] S. Partridge (1962). Elastin. *Advances in Protein Chemistry,* **17**, 227–302.

[14] Li, B., Alonso, D., & Daggett, V. (2001). The molecular basis for the inverse temperature transition of elastin. *Journal of Molecular Biology,* **305**, 581–592. doi:10.1006/jmbi.2000.4306.

[15] Hoeve, C. & Flory, P. (1974). The elastic properties of elastin. *Biopolymers,* **13**, 677–686.

[16] Wasserman, Z. & Salemme, F. (1990). A molecular dynamics investigation of the elastomeric restoring force in elastin. *Biopolymers,* **29**, 1613–1631.

[17] Li, B. & Daggett, V. (2002). Molecular basis for the extensibility of elastin. *Journal of Muscle Research and Cell Motility,* **23**, 561–573.

[18] Fleming, W., Sullivan, C., & Torchia, D. (1980). Characterization of molecular motions in ^{13}C-labeled aortic elastin by [13]C-^{1}H magnetic double resonance. *Biopolymers,* **19**, 597–617.

[19] Torchia, D. & Piez, K. (1973). Mobility of elastin chains as determined by ^{13}C nuclear magnetic resonance. *Journal of Molecular Biology,* **76**, 419–424.

[20] Aaron, B. & Gosline, J. (1980). Optical properties of single elastin fibres indicate random protein conformation. *Nature,* **287**, 865–867.

[21] Urry, D., Trapane, T., Long, M., & Prasad, K. (1983). Test of the librational entropy mechanism of elasticity of polypentapeptide of elastin. *J.Chem.Soc., Faraday Trans. 1,* **79**, 853–868.

[22] O. Arad, M. G. (1990). Depsipeptide analogues of elastin repeating sequences: conformational analysis. *Biopolymers,* **29**, 1651–1668.

[23] Urry, D., Shaw, R., & Prasad, K. (1985). Polypentapeptide of elastin: temperature dependence of ellipticity and correlation with elastomeric force. *Biochemical and Biophysical Research Communications,* **130**, 50–57.

[24] Schreiner, E., Nicolini, C., Ludolph, B., Ravindra, R., Otte, N., Kohlmeyer, A., Rousseau, R., Winter, R., & Marx, D. (2003). Folding and unfolding of an elastin–like oligopeptide: "inverse temperature transition", re-entrance, and hydrogen–bond dynamics. *submitted,* .

[25] Nicolini, C., Ravindra, R., Ludolph, B., & Winter, R. (2003). Characterization of the temperature- and pressure-induced inverse and re-entrant transition of the minimum elastin-like polypeptide GVG(VPGVG) by DSC, PPC, CD and FT-IR spectroscopy. *Journal of Molecular Biology,* **0**, 0–0. previous paper.

[26] Debelle, L., Alix, A., Jacob, M., Huvenne, J., Berjot, M., Sombret, B., & Legrand, P. (1995). Bovine elastin and kappa-elastin secondary structure determination by optical spectroscopies. *Journal of Biological Chemistry,* **270**, 26099–26103.

[27] Perry, A., Stypa, M., Foster, J., & Kumashiro, K. (2002). Observation of the glycines in elastin using ^{13}C and ^{15}N solid-state NMR spectroscopy and isotopic labeling. *Journal of American Chemical Society,* **124**, 6832–6833. doi:10.1021/ja017711x.

[28] Kurková, D., Kříž, J., Schmidt, P., Dybal, J., Rodríguez-Cabello, J., & Alonso, M. (2003). Structure and dynamics of two elastin-like polypentapeptides studied by NMR spectroscopy. *Biomacromolecules,* **4**, 589–601. doi:10.1021/bm025618a.

[29] Li, B., Alonso, D., Benion, B., & Daggett, V. (2001). Hydrophobic hydration is an important source of elasticity in elastin-based biopolymers. *Journal of American Chemical Society,* **123**, 11991–1198. doi:10.1021/ja010363e.

[30] Li, B., Alonso, D., & Daggett, V. (2002). Stabilization of globular proteins via introduction of temperature-activated elastin-based switches. *Structure,* **10**, 989–998.

[31] Urry, D., Trapane, T., Iqbal, M., Venkatachalam, C., & Prasad, K. (1985). Carbon-13 NMR relaxation studies demonstrate an inverse temperature transition in the elastin polypentapeptide. *Biochimie,* **24**, 5182–5189.

[32] Savitzky, A. & Golay, M. (1964). Smoothing and differentiation of data by simplified least squares procedures.

Analytical Chemistry, **36**, 1627–1639.

[33] Dixit, S., Crain, J., Poon, W., Finney, J., & Soper, A. (2002). Molecular segregation observed in a concentrated alcohol-water solution. *Nature,* **416**, 829. doi:10.1038/416829a.

[34] García, A. E. (1992). Large-amplitude nonlinear motions in proteins. *Physical Review Letters,* **68**, 2696–2699. doi:10.1103/PhysRevLett.68.2696.

[35] Amadei, A., Linssen, A., & Berendsen, H. (1993). Essential dynamics of proteins. *Proteins: Structure, Function and Genetics,* **17**, 412–425.

[36] Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis* chapter 6. John Wiley & Sons.

[37] Schlitter, J. (1993). Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chemical Physics Letters, ,* **215**, 617–621. doi:10.1016/0009-2614(93)89366-P.

[38] Andriciolaei, I. & Karplus, M. (2001). On the calculation of entropy from covariance matrices of the atomic fluctuations. *Journal of Chemical Physics, ,* **115** (14), 6289–6292. doi:10.1063/1.1401821.

[39] Stillinger, F. H. (1975). Theory and molecular models for water. *Advances in Chemical Physics, ,* **31**, 1–101.

[40] Luzar, A. & Chandler, D. (1996). Effect of environment on hydrogen bond dynamics in liquid water. *Physical Review Letters, ,* **76**, 928–931. doi:10.1103/PhysRevLett.76.928.

[41] Luzar, A. & Chandler, D. (1996). Hydrogen bond kinetics in liquid water. *Nature, ,* **379**, 55–57.

[42] Starr, F., Nielsen, J., & Stanley, H. (1999). Fast and slow dynamics of hydrogen bonds in liquid water. *Physical Review Letters, ,* **82**, 2294–2297. doi:10.1103/PhysRevLett.82.2294.

[43] Chandra, A. (2000). Effects of ion atmosphere on hydrogen-bond dynamics in aqueous electrolyte solutions. *Physical Review Letters, ,* **85**, 768–771. doi:10.1103/PhysRevLett.85.768.

[44] Xu, H. & Berne, B. (2001). Hydrogen-bond kinetics in the solvation shell of a polypeptide. *Journal of Physical Chemistry B, ,* **105**, 11929–11932. doi:10.1021/jp012749h.

[45] Tarek, M. & Tobias, D. (1999). Environment dependence of the dynamics of protein hydration water. *Journal of American Chemical Society, ,* **121**, 9740–9741. doi:10.1021/ja990643i.

[46] Tarek, M. & Tobias, D. (2002). Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Physical Review Letters, ,* **88**, 138101–1–138101–4. doi:10.1103/PhysRevLett.88.138101.

[47] Tarek, M. & Tobias, D. (2002). Single-particle and collective dynamics of protein hydration water: a molecular dynamics study. *Physical Review Letters, ,* **89**, 275501–1–275501–4. doi:10.1103/PhysRevLett.89.275501.

[48] Luzar, A. (2000). Extent of inter-hydrogen bond correlations in water. temperature effect. *Chemical Physics, ,* **258**, 267–276.

[49] van der Spoel, D., van Maaren, P., & Berendsen, J. (1998). A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *Journal of Chemical Physics, ,* **108**, 10220–10230.

[50] Brooks III, C. & Karplus, M. (1983). Deformable stochastic boundaries in molecular dynamics. *Journal of Chemical Physics, ,* **79**, 6312–6325.

[51] MacKerell Jr., A. D. *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B, ,* **102**, 3586–3616.

[52] Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R., & Klein, M. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics, ,* **79**, 926–935.

[53] Eichinger, M., Heller, H., & Grubmüller, H. (2000). EGO – an efficient molecular dynamics program and its application to protein dynamics simulations. In: *Molecular Dynamics on Parallel Computers,* (Esser, R., Grassberger, P., Grotendorst, J., & Lewerenz, M., eds), pp. 154–174, Singapore: World Scientific.

[54] Eichinger, M., Grubmüller, H., Heller, H., & Tavan, P. (1997). FAMUSAMM: An algorithm for rapid evaluation of electrostatic interactions in molecular dynamics simulations. *Journal of Computational Chemistry,* **18**, 1729–1749.

[55] Tamura, T., Yamaoka, T., Panitch, A., & Tirrell, D. (2000). Effects of temperature and pressure on the aggregation properties of an engineered elastin model polypeptide in aqueous solution. *Biomacromolecules,* **1**, 552–555. doi:10.1021/bm005606u.

[56] McLachlan, A. D. (1979). Gene duplications in the structural evolution of chymotrypsin. *Journal of Molecular Biology,* **128**, 49–77.

[57] Aalten, D. V., Groot, B. D., Findlay, J., Berendsen, H., & Amadei, A. (1997). A comparison of techniques for calculating protein essential dynamics. *Journal of Computational Chemistry,* **18**, 169–181. doi:10.1002/(SICI)1096-987X(19970130)18:2<169::AID-JCC3>3.0.CO;2-T.